几何距优化质心结合隶属度约束 RFCM 的脑 MRI 图像分割算法 *

南丽丽 [1]，邓小英 [2]

(1. 运城学院 计算机科学与技术系，山西 运城 044000; 2. 北京理工大学 信息与电子学院，北京 100081)

摘　要：针对现有图像分割算法聚类复杂以及分割精度不够高的问题，提出了基于几何距优化质心和粗糙模糊 C-均值(RFCM)相结合的医学图像聚类分割算法。首先建立软集表示的像素集，并计算每个像素与质心之间的距离，然后基于像素和质心之间的最小距离，将像素分组到聚类中。为了将软集应用到粗糙模糊 C-均值中，定义了一个模糊软集，进一步将输入图像转换为二值图像，通过计算连通区域的几何距选择适当的质心。最后利用这些新的质心计算更新像素的隶属度值，从而完成模糊聚类划分。在 Allen Brain Atlas 等三个医学数据库上评估了所提出混合算法的性能，获得的 Jaccards 系数和分割精度(SA)都优于几种对比算法。实验证明，提出的聚类分割算法具有良好的性能。

关键词：图像分割；软集合理论；质心计算；粗糙区域；模糊 C-均值；隶属度约束

中图分类号：TP391.41　　doi: 10.3969/j.issn.1001-3695.2018.07.0428

RFCM magnetic resonance brain image segmentation algorithm based on geometric distance optimization of centroid and membership constraints

Nan Lili[1], Deng Xiaoying[2]

(1. *Dept. of Computer Science & Technology, Yuncheng University, Yuncheng Shanxi* 044000, *China*; 2. *School of Information & Electronics, Beijing Institute of Technology, Beijing* 100081, *China*)

Abstract: Aiming at the problem that the existing image segmentation algorithm is complex and the segmentation precision is not high enough, proposed a medical image clustering segmentation algorithm based on geometric distance optimization centroid and rough fuzzy C-means (RFCM) . First set up the set of pixels represented by the soft set and calculate the distance between each pixel and the centroid. The pixels are then grouped into clusters based on the minimum distance between the pixel and the centroid. In order to apply the soft set to the coarse fuzzy C-means, a fuzzy soft set is defined. The input image is further converted into a binary image, and an appropriate centroid is selected by calculating the geometric distance of the connected region. Finally, using these new centroids to calculate the membership value of the updated pixel, the fuzzy clustering is completed. The performance of the proposed hybrid algorithm was evaluated on three medical databases, such as Allen Brain Atlas. The obtained Jaccards coefficient and segmentation accuracy (SA) were better than several comparison algorithms. Experiments show that the proposed clustering segmentation algorithm has good performance.

Key words: image segmentation; soft set theory; centroid calculation; rough area; fuzzy C-means; membership constraint

0　引言

图像分割是图像分析和处理中一项关键的功能，在描绘解剖学、医学图像诊断、组织液提取、病理学局部化、治疗计划、功能图像数据的部分容积校正以及计算机集成外科手术等方面有着巨大的应用[1-3]。模糊集、粗糙集是处理医学图像不确定性和模糊性的有效工具，被广泛应用于医学图像分割[4]。软集合理论是处理不确定性和模糊性的一种新的数学方法[5]。

经典的 K-均值聚类算法或者硬 C-均值聚类算法都是基于清晰集[6]，其中每一个像素只属于一个聚类。它的模糊版本：模糊 C-均值聚类算法[7]也受到了极大的关注。模糊 C-均值方法没有考虑空间条件和边界条件，因此对噪声和不充分数据很敏感。为了考虑空间条件和边界区域，一些研究开始把软集合理论纳入到 KM 和 FCM 框架中进行图像分割[5]。还有一些文献中提出了混合算法：粗糙 K 均值[8]、粗糙模糊可能性 C-均值[9]、阴影 C-均值[10]等。文献[12]提出了一个粗糙集的算法，使用一个阈值而不是负域，来找到下近似和上近似。然而，由于像素绝对属于下近似和上近似，或者两者都有，因此聚类时不会消除像素。文献[13]提出了软粗糙集，用以定义不包含负域的下近似和上近似。但是这种技术使用聚类原型的随机初始化，当

收稿日期：2018-07-12；**修回日期：**2018-08-30　　**基金项目：**国家自然科学基金资助项目（41374114）

作者简介：南丽丽（1978-），女，山西运城人，副教授，硕士，主要研究方向为图像处理（xie531596jia2272@163.com）；邓小英（1974-），女，四川广安人，副教授，博士，硕导，主要研究方向为数字图像处理、数字信号处理.

选择错误时会产生不准确的结果。文献[14]提出了一种计算模糊软集的下近似和上近似的方法，以有效地定义边界。然而，这些定义是针对清晰值，并不能处理边界像素，粗糙区域的确定也异常复杂。这几种基于粗糙集的算法，都涉及复杂的指令来计算近似值，并且基于单一阈值将像素分成粗糙区域。试图依据聚类的初始质心，使用像素的最小和最大隶属度[11]来计算阈值。因此，如果选择了不恰当的初始质心，会导致不准确的结果。此外，在文献[15]中提出了将直觉模糊集和粗糙集相结合的混合算法。这些算法使用了三元隶属度，表示从属于一个聚类、不属于一个聚类、不确定是否属于一个聚类。使用三元隶属度来计算上下近似和区域。因此，这些算法涉及的计算比粗糙集更复杂。

本文提出了使用软集表示图像的观点，并用软模糊粗糙近似定义了图像粗糙区域的下近似和上近似。与粗糙集算法不同，这种聚类算法避免了阈值、权重参数以及复杂上下近似的使用。该算法不使用负区域，因此所有的像素都可以参与聚类。软集利用相似性度量来量化不确定性，这个不确定性跟前一个聚类和当前聚类的相似程度有关。

1　粗糙集理论

粗糙集对不精确概念的描述方法是通过下近似和上近似概念来描述。使用粗糙集下近似和上近似来近似不确定数据。设 U 表示全集，R 是等价关系，U/R 是 n 个等价类 $\{x_1,x_2,\cdots,x_n\}$ 的集合，它们在 X 中形成分区。(U,R) 是近似空间。对子集 $X\subseteq U$，下近似和上近似可以表示为

$$\begin{aligned}\underline{R}(x_i)&=\bigcup_{x_i\in X}x_i\\ \overline{R}(x_i)&=\bigcup_{x_i\cap X\neq\phi}x_i\end{aligned}\tag{1}$$

$\underline{R}(x_i)$ 表示对象 x_i 属于 X 时，X 的下近似空间。$\overline{R}(x_i)$ 表示对象可能属于 X 时，X 的上近似空间。X 的近似空间被分类为三个不同的区域：

$$\begin{aligned}\text{正域}(x_i)&=\underline{R}(x_i)\\ \text{边界域}(x_i)&=\overline{R}(x_i)-\underline{R}(x_i)\\ \text{负域}(x_i)&=U-\overline{R}(x_i)\end{aligned}\tag{2}$$

一些研究中利用粗糙集理论，提出了一些混合算法，如：粗糙 K-均值算法和粗糙模糊 C-均值算法。如果一个像素属于所有的聚类，则其隶属度较少。把该像素包括到聚类中，会产生噪声；从各聚类中去除该像素，会产生不足的数据。聚类中的这些不完整性、不确定性和模糊性可以使用粗糙集的下近似和上近似来定义。聚类的下近似是绝对属于聚类的像素集合，而上近似是可能属于聚类的像素集合。给予边界区域（上下近似之间）中的像素第二次机会，以便在下一次迭代中它们可以移动到合适的聚类。因此，粗糙集可以减少聚类错误。

2　粗糙模糊 C-均值聚类

模糊 C-均值聚类算法是从硬 C-均值聚类算法发展而来，它采用迭代法优化目标函数来获取对数据集的模糊分类，算法具有很好的收敛性，是实际应用最为广泛的一种模糊聚类算法。

文献[16]提出一种用于磁共振脑图像的模糊 C-均值聚类方法，其偏置场模型的定义如方程（3）所示。以下等式给出了分割磁共振脑图像的目标函数：

$$\begin{aligned}J_{GFCM}&=\sum_{j=1}^{k}\left(W_l\sum_{x_i\in P(C_j)}\left\|x_i-b_iC_j\right\|^2\right)\\&+\sum_{j=1}^{k}\left((1-W_l)\sum_{x_i\in B(C_j)}\left(u_{ij}\right)^m\left\|x_i-b_iC_j\right\|^2\right)\end{aligned}\tag{3}$$

其中：u_{ij} 隶属度值是用下面的带偏置场的等式计算[17]。

$$u_{ij}=\frac{1}{\sum_{c=1}^{k}\left(\frac{\left\|x_i-b_iC_j\right\|}{\left\|x_i-b_iC_c\right\|}\right)^{\frac{2}{m}-1}}\tag{4}$$

根据两个阈值 t_1 和 t_2 计算正域 $P(C_j)$ 和边界域 $B(C_j)$。这两个阈值定义如下：

$$\begin{aligned}t_1&=\frac{1}{n}\sum_{i=1}^{n}d_{\min}\\ t_2&=\frac{1}{n}\sum_{i=1}^{n}d_{\max}\end{aligned}\tag{5}$$

其中：$d_{\min}$ 是像素与质心的最小距离，$d_{\max}$ 是像素与质心的最大距离。

$$x_i=\begin{cases}P(C_j);\ \text{若}d_{ij}\le t_1\\ B(C_j);\ \text{若}t_1\le d_{ij}\le t_2\\ N(C_j);\ \text{其他}\end{cases}\tag{6}$$

其中:每个聚类的权重参数被估计为正域与整个区域的百分比，由下式给出：

$$W_l=\frac{\left|P(C_j)\right|}{x_i=\arg\max\left(u_{ij};i=1,\dots,n;j=1,\dots,k\right)}\tag{7}$$

距离 d_{ij} 由下式给出：

$$d_{ij}=\arg\min\left(\left\|x_i-C_j\right\|\right)^2\tag{8}$$

通过正交多项式的线性组合来逐个像素地近似偏置场 b。

$$b_i=\sum_{i}^{M}w_kg_k(i)=W^TG(i)\tag{9}$$

其中：$M=(D+1)(D+1)$ 是度为 D 的多项式 $g_k(i)$ 的数量，而 $W^TG(i)$ 是基函数集合 $G(i)$ 的线性组合。使用能量函数估计和校正区域 Ω 中像素 i 的偏置场，如下式所示：

$$F=\sum_{j}^{k}\sum_{i\in\Omega}\left\|x_i-W^TG(i)C_j\right\|^2u_{ij}^m \quad (10)$$

能量相对于 W 最小化；可以通过求解 $DF/DW=0$ 来完成这一步，得到：

$$\frac{dF}{dW}=-2V+2A \quad (11)$$

其中: A 是由下式给出的多维列向量：

$$A=\sum_{i\in\Omega}G(i)G(i)^T\sum_{j=1}^{k}u_{ij}^mC_j^2 \quad (12)$$

$$V=\sum_{i\in\Omega}x_iG(i)^T\sum_{j=1}^{k}u_{ij}^mC_j \quad (13)$$

方程 $DF/DW=0$ 表示为一个线性方程：$AW=V$。通过解 $W=A^{-1}V$，偏置场可以估计为 $b_i=W^TG(i)$。在这种方法中，由阈值 t_1 和 t_2 来定义粗糙区域，距离大于 t_2 的像素落入负域，因此从聚类中消除该像素。

3 提出的方法

3.1.1 粗糙模糊 C-均值聚类

基于模糊的聚类允许重叠聚类。为了把数据 $\{x_1,x_2,\cdots,x_n\}$ 分割进模糊聚类 $\{c_1,c_2,\cdots,c_k\}$，为提高 FCM 算法的抗噪性，我们在 FCM 的目标函数中对隶属度引入空间约束关系，新的目标函数为

$$J=(U,\ v)=\sum_{k=1}^{C}\sum_{i,j}u_k^m(i,j)d_k^2(i,j)+\gamma\sum_{k=1}^{C}\left(\sum_{i',j'\in N_{i,j}}\left\|u_k(i,j)-u_k(i',j')\right\|\right) \quad (14)$$

其中：$N_{i,j}$ 为像素点的邻域区域，上式的第二项对 (i,j) 点隶属度 $u_k(i,j)$ 与其领域隶属度 $u_k(i',j')$ 的差异进行惩罚。为求解隶属度函数，建立 Lagrange 函数，可得

$$F=\sum_{k=1}^{C}\sum_{i,j}u_k^m(i,j)d_k^2(i,j)+\lambda_{i,j}\left(1-\sum_{k=1}^{C}(i,j)\right)+\gamma\sum_{k=1}^{C}\left(\sum_{i',j'\in N_{i,j}}\left\|u_k(i,j)-u_k(i',j')\right\|\right) \quad (15)$$

u_{ij} 表示第 i 个像素对第 j 个模糊聚类的隶属程度，由式（16）给出。

$$u_{ij}=\frac{1}{\sum_{c=1}^{k}\left(\frac{\left\|x_i-C_j\right\|}{\left\|x_i-C_c\right\|}\right)^{\frac{2}{m}-1}} \quad (16)$$

C_j 是聚类 j 的平均值，m 是用户定义的模糊值。如果模糊值 m 接近于 1，隶属度函数变得更加清晰明了，并且如果 m 增加，隶属度函数变得越来越模糊不清。对于大部分数据，m 的值为 1.5> m <3，能够获得很好的结果，通常取 2。W_l 和 $W_h=1-W_l$ 是权重因子，影响下近似和边界区域的相对重要性。每个聚类都与下近似 $\underline{R}(C_j)$、上近似 $\overline{R}(C_j)$ 和边界区域 $B(C_j)=\overline{R}(C_j)-\underline{R}(C_j)$ 相关联。X 中的所有元素都是通过使用一个阈值来分组到下近似区域 $\underline{R}(C_j)$ 或边界区域 $B(C_j)$ 或上近似区域 $\overline{R}(C_j)$ 中，该阈值是使用元素 x_i 到聚类质心 C_j 的最高隶属度 u_{ij} 来计算的。粗糙模糊 C-均值使用式（17）中定义的阈值，将像素置于下近似或上近似中。

$$T=\left(\frac{u_{ij}}{\arg\max(u_{ij})}\right)\le\text{阈值} \quad (17)$$

如果 $T\ne\phi$，那么 $x_i\in\overline{R}(C_j)$ 并且 $x_i\in\overline{R}(C_k)$，否则 $x_i\in\overline{R}(C_j)$ 并且 $x_i\in\underline{R}(C_j)$。如果像素满足阈值条件，那么它们将被放置在下近似中，否则它们将被放置在边界区域中或者被给予第二次机会。然而，由于缺乏参数化工具，使用粗糙集计算近似值是比较复杂的。而且，阈值和权重参数是用户定义的值。

3.1.2 模糊软集

软集是将参数映射到全集。令 U 为全集，A 是参数集，使得 $A\subseteq E$。那么 U 上的软集由 (F,A) 给出，其中 F 是一个映射：F：$A\to P(U)$。$P(U)$ 表示 U 的幂集，包含 U 的所有子集[28]。把软集应用于图像分割首先把观察图像中的像素集表示为 U，$U=X=\{x_i|x_i\text{是图像中第}i\text{个像素的值}\}$。假定观察图像 X 是由真实信号通过偏置场 b 和零均值高斯噪声得到的。然后设 (F,A) 是将要构建的软集，包含属于聚类的像素。最后把参数集作为像素所属聚类 $C\{i=1,2,3,...,k\}$ 数量。假设图像中的像素集由 $U=\{x_1,x_2,x_3,x_4,x_5\}$ 给出，参数集 A 为像素所属聚类 $\{C_1,C_2,C_3\}$。计算每个像素与质心之间的距离。基于像素和质心之间的最小距离，按如下方法将像素分组到聚类中。

$$\begin{aligned}F(C_1)&=\{x_1,x_6\}\\F(C_2)&=\{x_1,x_2,x_4,x_5\}\\F(C_3)&=\{x_3,x_4,x_5,x_6\}\end{aligned} \quad (18)$$

上面例子的软集表示为

$$(F,A)=\begin{Bmatrix} C_1=\{x_1,x_6\} \\ C_2=\{x_1,x_2,x_4,x_5\} \\ C_3=\{x_3,x_4,x_5,x_6\} \end{Bmatrix} \quad (19)$$

表 1 为基于软集表示的像素集。如果 $x_i \in F(C_j)$，那么其值为 1，否则就是 0。

表 1　基于软集表示的像素集

U	C_1	C_2	C_3
x_1	0	1	1
x_2	1	0	0
x_3	0	1	1
x_4	0	1	0
x_5	1	0	1
x_6	1	1	1

为了将软集应用到模糊值中，定义了一个模糊软集：模糊软集 (F,A) 是 U 的模糊值子集的参数簇。因此，(F,A) 可以表示为 $(F,A)=\{F(\ell)\,|\,\ell\in A\}$。$\{F(\ell)\}$ 表示 U 的模糊子集，它是 E 的模糊值集合参数：$F(\ell)=\{(x,\mu_f)\}$。因此，x 属于 U，其隶属度 $\mu_f(x)$ 为参数。因此，通过考虑每个像素对聚类质心的模糊隶属度，以软模糊集表示图像。模糊软集是参数对具有模糊值的全集映射。因此，为了表示模糊软集，使用等式（20）计算像素的隶属度值。假设像素对聚类的隶属度值如下：

$$\begin{aligned} F(C_1)&=\{(x_1,0.6),(x_2,0.1),(x_3,0.4),(x_4,0.7)\} \\ F(C_2)&=\{(x_1,0.4),(x_2,0.2),(x_3,0.5),(x_4,0.1)\} \\ F(C_3)&=\{(x_1,0.4),(x_2,0.3),(x_3,0.5),(x_4,0.2)\} \end{aligned} \quad (20)$$

另外，模糊集 $\mu=\{(C_1,0.6),(C_2,0.6),(C_3,0.6)\}$，这样一个 0.6 隶属度的像素能被放置在聚类中。选择 0.6 的值是为了包括在聚类中的像素至少有 0.6 的隶属度值。如果它小于 0.5，那么聚类会变得更加模糊并增加噪声。关于软集的 μ 模糊粗糙的下近似和上近似由 $\underline{R}\mu(x)$ 和 $\overline{R}\mu(x)$ 定义。

$$\underline{R}\mu(x)=\Lambda_{\ell\in A}\left(1-\left(f(\ell)(x)\right)\vee\left(\Lambda_{y\in U}\left(\left(1-F(\ell)(y)\right)\vee\mu(y)\right)\right)\right) \quad (21)$$

$$\overline{R}\mu(x)=\Lambda_{\ell\in A}\left(\left(f(\ell)(x)\right)\wedge\left(\vee_{y\in U}\left(\left(F(\ell)(y)\right)\wedge\mu(y)\right)\right)\right) \quad (22)$$

把 U 作为 X，μ 作为模糊集，U_{ij} 作为 $F(1)$，方程（21）和（22）可以修改成如下形式以用于图像分割。

$$\underline{R}\mu(x)=\Lambda_{x_i\in C_j}\left(\left(1-\mu_{ij}\right)\vee\left(\Lambda_{x_i\in X}\left(1-\mu_{ij}\right)\vee\mu(j)\right)\right) \quad (23)$$

$$\overline{R}\mu(x)=\Lambda_{x_i\in C_j}\left(\left(\mu_{ij}\right)\vee\left(\vee_{x_i\in X}\left(\mu_{ij}\right)\wedge\mu(j)\right)\right) \quad (24)$$

使用等式（23）和（24）的模糊集 μ 计算软模糊集的下近似和上近似，下近似和上近似的表格表示在表 2 中给出。

表 2　下近似值和上近似值

U	C_1	C_2	C_3	$\underline{R}\mu(x)$	$\overline{R}\mu(x)$
x_1	0.5	0.6	0.4	0.4	0.6
x_2	0.1	0.2	0.2	0.7	0.1
x_3	0.4	0.6	0.5	0.4	0.4
x_4	0.6	0.1	0.2	0.5	0.2
μ	0.6	0.7	0.6	—	—

在上面的定义中，没有使用阈值或权重参数来计算近似值，而是使用了简单的与、或操作。使用这些近似值，就可以通过以下等式确定正域 $P(C_j)$ 和边界域 $B(C_j)$。

$$x_i=\begin{cases} P(C_j)；\ 若 x_i\ge\underline{R}\mu(x_i) \\ B(C_j)；\ 若 \overline{R}\mu(x_i)\ge x_i\le\underline{R}\mu(x_i) \end{cases} \quad (25)$$

3.1.3 几何距离优化质心

如果一个像素的度数在 $\overline{R}\mu(x)$ 内，那么把该像素置于边界域；否则，如果该像素具有比 $\underline{R}\mu(x)$ 更大的值，则把该像素置于正域中。一旦获得了边界域和正域，那么就可以基于正域和负域，使用下面公式中的偏置场 b_i 来计算新的质心。

$$C_j=\frac{\sum_{x_i\in P(C_j)}x_ib_i}{\sum_{x_i\in P(C_j)}b_i^2}+\frac{\sum_{x_i\in B(C_j)}\left(u_{ij}\right)^m x_ib_i}{\sum_{x_i\in B(C_j)}\left(u_{ij}\right)^m b_i^2} \quad (26)$$

利用这些新的质心，计算更新像素的隶属度值。由于聚类算法是迭代的，基于软集的聚类算法使用前一次迭代和当前迭代形成的聚类之间的相似度进行迭代。

初始聚类质心在分割中扮演着重要的角色，在初始隶属度函数中使用如 K-均值聚类算法或者硬 C 均值聚类算法这样的聚类算法。错误选择的质心可能导致局部最小从而导致不准确的结果。将输入图像转换为二值图像，并选择相应的灰度作为初始聚类质心。几何距优化质心的算法步骤如下：

a)将输入图像转换为二值图像；

b)通过连通组件标记算法找到所有的连通区域，并分别标记；

c)对每个连通区域运用计算几何距算法得到质心；

$$M_{pq}=\int_{a_1}^{a_2}\int_{b_1}^{b_2}x^py^qf(x,y)dxdy \quad (27)$$

d)再用不同颜色绘制出连通区域与质心，进一步输出处理后图像；

软模糊粗糙集是通过嵌入软集、粗糙集和模糊集的概念得到的。软粗糙集使用表示归属程度的软模糊集来定义聚类的下近似和上近似。每个像素的近似值都是用模糊值而不是清晰值来定义的。

4　实验结果

4.1　方法实施

在三个医学数据库上评估了所提出混合算法的性能：（i）脑图像数据库 Brain Perfusion Database[18]；（ii）艾伦大脑图谱 Allen Brain Atlas[19]；（iii）BRATS 数据库[20]。这些数据库中选取的供仿真所用图像的数量以及大小如表 3 所示。这三个数据库都是经典的脑图像数据库，图像的数量以及大小都适合本文方法进行仿真实验。把这些图像转换成 MATLAB 可读的格式，作为预处理步骤，在进行分割之前，使用 MIPAV（医学图像处理、分析和可视化）软件将非脑部图像从图像中移除。所提出的算法在 4.2GHz Intel I7 7700K 处理器上的 MATLAB 2012a 中实现。图 1 为艾伦大脑图谱中的部分图像。



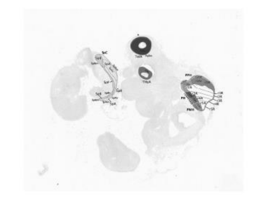
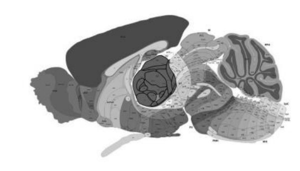



图 1　艾伦大脑图谱中的部分图像

表 3　仿真所用数据集

维度	数据集	数量	图像大小
2D	Brain Perfusion Database	10	194*237
2D	Allen Brain Atlas	10	256*128
2D	BRATS	10	变量*216

磁共振脑图像经常分为白质（WM）、灰质（GM）和脑脊液（CSF）。有效提取这些区域对于定量分析是非常重要的，因此，产生了分割 MR 脑图像的几种方法。磁共振脑图像如图 2 所示。

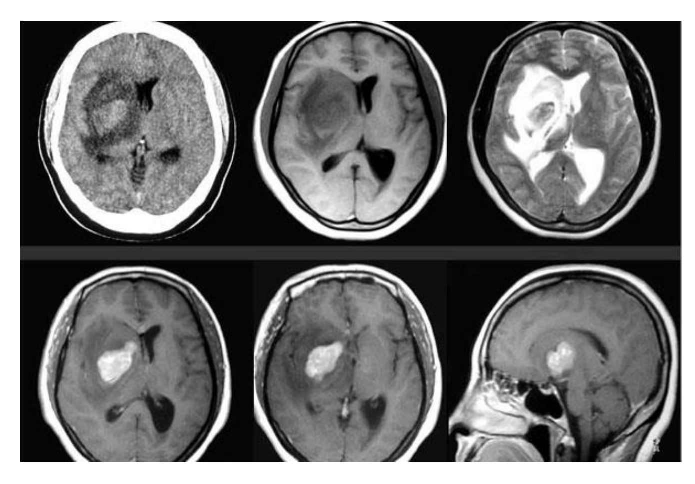


图 2　磁共振脑图像

在所有大脑图像的实验设置中，取聚类数量 k =4（关于 GM、WM、CSF 和背景）。使用与实际分割边界的定量比较来计算分割结果，使用 Jaccards 系数(Jaccard Coefficient, JC)和分割精度(Segmentation Accuracy, SA)对分割结果进行评估。两个数据集之间的 JC 是两个数据集共有的属性数量除以所有属性数量的结果。

$$Jaccard\text{系数} = \frac{A \cap B}{A \cup B} \tag{29}$$

其中: A 是本文方法获得的分割图像， B 是真实图像。Jaccard 系数在 70%以上，这意味着分割结果是好的。SA 是用来比较分割图像和真实图像之间相似性的另一个常用指标。为了计算分割精度，对每个分割结果计算四个参数。

a)真正(TP)：真实图像中正确像素被正确检测为分割像素。

b)真负(TN)：真实图像中错误像素被正确检测为分割像素。

c)假正(FP)：分割区域内未发现真实图像中的正确像素。

d)假负(FN)：分割区域中不存在真实图像中的错误像素。

$$SA = \frac{TP + FN}{TP + FP + TN + FN} \tag{30}$$

第一个实验在脑图像数据库 Brain Perfusion Database 数据库上进行，该数据库包含 T1 加权仿真大脑 10 张图像的二维轴向视图，片数为 90，厚度为 1mm，不同强度不均匀性（INU = 0 和 INU = 20）以及噪声级别（0%、1%、3%、5%和 9%）。由于磁共振脑图像数据的典型噪声级别在 3%左右，因此选择含有 3%噪声和 0%INU 的合成脑磁共振图像。为了验证实验，还选择了其他噪声级别为 9%的噪声图像。

4.2　性能评估

许多聚类算法的性能受初始聚类质心的影响。本文提出的直方图优化质心的方法解决了关于随机质心的问题。基于直方图的优化质心在寻找接近真实质心的起始点方面实现了高效率，从而得到可靠的脑图像分割结果。表 4 给出了算法的 SA 结果，表 5 中给出了算法的 JC 值。结果表明，即使存在噪声，提出的新算法也表现得非常好。

文献[12]方法虽然花费很少的时间进行评估，但对存在于边界区域（不可辨认）的像素却不起作用，因此增加了聚类错误，降低了 SA 和 JC 值。虽然通过定义隶属度，K 均值聚类算法或者硬 C-均值聚类算法方法能够来处理模糊性，但对不可分辨的情况却不能恰当地聚类。因此，一个组织的 SA 和 JC 值增加，另一个组织的 SA 和 JC 值就会减小。通过提供充当缓冲区域的边界区域，粗糙集可以有效地处理聚类中的不可分辨像素。在分配到聚类之前，给予缓冲区中的像素第二次机会，以减少聚类错误，获得良好的 SA 和 JC 结果，如表 4 表 5 以及图 3 所示。而文献[13-15]方法基于单一阈值将像素分成粗糙区域。试图依据聚类的初始质心，使用像素的最小和最大隶属度来计算阈值。因此，如果选择了不恰当的初始质心，会导致不准确的结果。

表 4　提出的方法与不同分割算法在不同噪声级别脑图像上仿真的性能比较(分割精度)

高斯噪声	组织类型	文献[12]	文献[13]	文献[14]	文献[15]	提出的方法
	GM	0.8123	0.8256	0.8055	0.8427	0.8976
3%	CSF	0.8352	0.8692	0.8460	0.8583	0.8892
	WM	0.8276	0.8027	0.8289	0.8167	0.9289
	GM	0.7967	0.7866	0.6268	0.6822	0.7992
6%	CSF	0.6976	0.8055	0.7199	0.7620	0.8027
	WM	0.6678	0.6311	0.7278	0.7856	0.7945
	GM	0.8027	0.7289	0.6678	0.6822	0.7978
9%	CSF	0.6918	0.7389	0.7278	0.7620	0.7892
	WM	0.6899	0.7027	0.7901	0.7856	0.8267

表 5　提出的方法与不同分割算法在不同噪声级别脑磁共振图像上仿真的性能比较(Jaccards 系数)

高斯噪声	组织类型	文献[12]	文献[13]	文献[14]	文献[15]	提出的方法
	GM	0.7826	0.8018	0.7822	0.7266	0.8729
3%	CSF	0.8036	0.8026	0.7811	0.8177	0.8677
	WM	0.8207	0.8237	0.8026	0.8027	0.8366
	GM	0.7826	0.7941	0.6622	0.6972	0.8027
6%	CSF	0.7027	0.7832	0.7598	0.7620	0.8062
	WM	0.6388	0.6432	0.7801	0.7893	0.7945
	GM	0.7372	0.7941	0.6748	0.6822	0.8072
9%	CSF	0.7026	0.7832	0.7528	0.7620	0.8053
	WM	0.6892	0.7027	0.7729	0.7836	0.8366

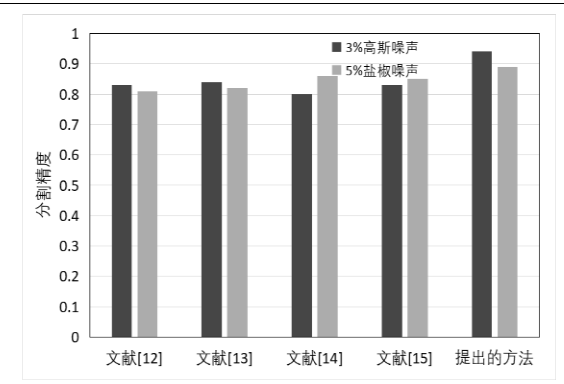


图 3　不同分割算法在不同噪声级别脑图像上仿真的性能比较(分割精度)

软集使用参数化工具有助于有效设计上近似和下近似。软集也使用相似性系数，将一个数量分配给代表相等程度的一对软集。与许多聚类算法类似，基于软集的算法也是迭代的，并且由于软集使用参数化工具和相似性系数，因此可以非常快速地达到收敛。为了处理不可分辨的像素并减少时间，粗糙集与软集一起使用被证明是更好的解决方案。对各个算法的执行时间进行比较，结果如图 4 所示。

由于使用参数化和相似性系数，所提出的方法避免了阈值的计算，从而降低了时间复杂度 $O(nk)$，其中 n 是图像大小、k 是聚类数量，而文献[12-15]方法都要计算阈值。因此，提出的方法比其他文献方法收敛更快。在计算量减少的同时，使用软集定义的磁共振脑图像粗糙区域保持了分割精度。

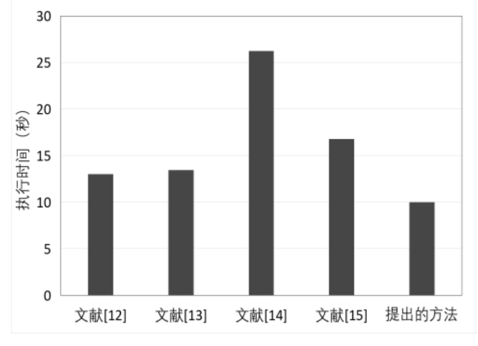


图 4　不同分割算法需要的执行时间比较

5　结束语

本文提出了基于粗糙模糊 C-均值隶属度约束的磁共振脑图像聚类分割算法。将软集应用于模糊集和粗糙集，提出了混合算法及其在医学图像分割中的应用。提出的算法具有许多优点：利用参数化工具定义上近似和下近似来处理不确定性，减少聚类错误，使用软集相似性系数达到快速收敛。该算法产生精确的结果。根据综合数据和仿真评估所提出算法的性能，所提出的方法提供了准确的大脑图像分割。未来希望能够对提出的方法进行进一步完善。

参考文献：

[1] 张海涛，程新文，熊红伟，等. 改进蜂群算法的图像阈值分割方法 [J]. 计算机应用研究，2017，34 (12): 3880-3884. (Zhang Haitao, Cheng Xinwen, Xiong Hongwei, *et al*. Image threshold segmentation method

based on improved artificial bee colony [J]. Application Research of Computers, 2017, 34 (12): 3880-3884.)

[2] 吕福起，李霄民．利用 MLT 和 SRS 的混合图像融合与去噪算法 [J]. 湘潭大学自然科学学报，2018, 40 (1): 111-114. (Lv Fuqi, Li Qiaomin. Hybrid image fusion and denoising algorithm using MLT and SRS [J]. Natural Science Journal of Xiangtan University, 2018, 40 (1): 111-114.)

[3] Fan Xian, Bazin P L, Bogovic J, *et al.* A multiple geometric deformable model framework for homeomorphic 3D medical image segmentation. [C]// Proc of IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops. 2016: 1-7.

[4] 马文萍，黄媛媛，李豪，等．基于粗糙集与差分免疫模糊聚类算法的图像分割 [J]. 软件学报，2014, 29 (11): 2675-2689. (Ma Wenping, Huang Yuanyuan, Li Hao, *et al.* Image segmentation based on rough set and differential immune fuzzy clustering algorithm [J]. Journal of Software, 2014, 29 (11): 2675-2689.)

[5] Rose A N M, Awang M I, Ahmad F, *et al.* Achieving efficient decision making through hybrid reduction in soft set theory [J]. 2017, 7 (3): 1032.

[6] 王勇，唐靖，饶勤菲，等．高效率的 K-means 最佳聚类数确定算法 [J]. 计算机应用，2014, 34 (5): 1331-1335. (Wang Yong, Tang Jing, Rao Qinfei, *et al.* High efficient K-means algorithm for determining optimal number of clusters [J]. Journal of Computer Applications, 2014, 34 (5): 1331-1335.)

[7] Azim S, Aggarwal S. Hybrid model for data imputation: Using fuzzy c means and multi layer perceptron [C]// Advance Computing Conference. IEEE, 2014: 1281-1285.

[8] 田大增，吴静．改进的粗糙模糊和模糊粗糙 K-均值聚类算法 [J]. 计算机工程与应用，2014, 50 (17): 142-145. (Tian Dazeng, Wu Jing. Improvement of rough fuzzy and fuzzy rough clustering algorithm [J]. Computer Engineering and Applications, 2014, 50 (17): 142-145.)

[9] 汪海良，余堃，周明天．基于阴影集的粗糙模糊可能性 C 均值聚类算法 [J]. 计算机科学，2013, 40 (1): 191-194. (Wang Hailiang, Yu Kun, Zhou Mingtian. Shadowed Sets-based Rough Fuzzy Possibilistic C-means Clustering [J]. Computer Science, 2013, 40 (1): 191-194.)

[10] 王丽娜，王建东，姜坚．特征加权的阴影 C-均值聚类新算法 [J]. 南京航空航天大学学报：英文版，2012, 29 (3): 273-283. (Wang Lina, Wang Jiandong, Jiang Jian. New shadowed C-means clustering with feature weights [J]. Transactions of Nanjing University of Aeronautics and Astronautics, 2012, 29 (3): 273-283.)

[11] Wang Jianqiang, Nie Rongrong, Zhang Hongyu, *et al.* New operators on triangular intuitionistic fuzzy numbers and their applications in system fault analysis [J]. Information Sciences, 2013, 251 (12): 79-95.

[12] Goswami S, Goswami S, Chakraborty S, *et al.* Image segmentation using rough set theory: a review [J]. International Journal of Rough Sets & Data Analysis, 2014, 1 (2): 62-74.

[13] Dubey Y K, Mushrif M M, Mitra K. Segmentation of brain MR images using rough set based intuitionistic fuzzy clustering [J]. Biocybernetics & Biomedical Engineering, 2016, 36 (2): 413-426.

[14] Abdullah S, Amin N U. Analysis of S-box image encryption based on generalized fuzzy soft expert set [J]. Nonlinear Dynamics, 2015, 79 (3): 1679-1692.

[15] Mukherjee A, Saha A, Das A. Soft sets combined with interval valued intuitionistic fuzzy sets of type-2 and rough sets [J]. New Trends in Mathematical Sciences, 2015, 3 (2): 199-218.

[16] Elazab A, Wang Changmiao, Jia Fucang, *et al.* Segmentation of brain tissues from magnetic resonance images using adaptively regularized kernel-based fuzzy C-means clustering [J]. Computational & Mathematical Methods in Medicine, 2015, 2015 (5): 485495.

[17] Tripathy B K, Tripathy A, Rajulu K G. Possibilistic rough fuzzy C-means algorithm in data clustering and image segmentation [C]// Proc of IEEE International Conference on Computational Intelligence and Computing Research. 2015: 1-6.

[18] Okita Y, Miyata H, Motomura N, *et al.* A study of brain protection during total arch replacement comparing antegrade cerebral perfusion versus hypothermic circulatory arrest, with or without retrograde cerebral perfusion: analysis based on the Japan Adult Cardiovascular Surgery Database [J]. Journal of Thoracic & Cardiovascular Surgery, 2015, 149 (2): 65-73.

[19] Hawrylycz M, Ng L, Feng David, *et al.* The Allen Brain Atlas [M]// Springer Handbook of Bio-//Neuroinformatics. Springer Berlin Heidelberg, 2014: 1111-1126.

[20] 侯发忠，邹北骥，刘召斌，等．基于灰度分布匹配的多模态脑部 MR 图像肿瘤分割算法 [J]. 计算机应用研究，2017, 34 (12): 3869-3872. (Hou Fazhong, Zou Beiji, Liu Zhaobin, *et al.* Tumor segmentation on multi-modality brain MR image based on grayscale distribution matching [J]. Application Research of Computers, 2017, 34 (12): 3869-3872.)